ChinaXiv合作期刊
图书情报工作
第64卷 第8期 2020年4月

# 国内外政府数据开放中的个人隐私保护研究述评*

■ 陈朝兵　郝文强
西南财经大学公共管理学院　成都 611130

摘　要：[目的/意义]系统梳理国内外政府数据开放中个人隐私保护的研究文献，旨在为国内该领域的进一步研究提供参考与借鉴。[方法/过程]运用文献调研法，同时采用归纳、比较、综合等分析方法，从政府数据开放中个人隐私保护的概念、缘由、逻辑悖论、实践困境、实现路径等方面进行综述。[结果/结论]结果表明，该领域研究总体经历了隐私议题出现、政府信息公开中的个人隐私保护和政府数据开放中的个人隐私保护3个阶段。相关主题的研究取得了较丰硕的成果，跨学科研究特征明显，研究观点既有共识也有分歧。建议国内研究从4个方面展开：拓展研究议题范畴，聚焦关键问题；把握研究问题的复杂性，注重应用系统思维；丰富理论工具的应用，突出研究的过程性和动态性；引入历史、比较等分析方法，加强案例实证研究。

关键词：政府数据开放　政府信息公开　个人隐私　隐私保护　文献述评
分类号：G203
DOI：10.13266/j.issn.0252-3116.2020.08.016

chinaXiv:202304.00260v1

自20世纪中叶人类社会踏入互联网与信息化时代伊始，个人隐私风险与安全问题便如影随形。进入21世纪大数据时代，席卷全球的政府数据开放运动在有力推动"数据民主化"（Democratizing Data）的同时，也进一步加剧了人们对隐私保护问题的隐忧。毋庸置疑，在世界各国由政府信息公开向政府数据开放"转型升级"的进程中，随着隐私泄露、隐私破坏、隐私侵犯等隐私安全问题的日益突出和严峻，隐私保护已成为各国政府和国际社会共同关注与高度重视的重大实践议题。

基于对现实的关注与回应，学术界早在19世纪80年代便开始了对个人隐私保护问题的研究，特别是近十年来在政府数据开放背景下投入了越来越多的注意力，从而积累了一大批富有价值的研究成果。那么，学术界对政府数据开放中个人隐私保护的研究经历了怎样的发展历程？开展了哪些工作？取得了何种成绩？存在哪些不足？等等。鉴于回答这一系列问题是推进该领域未来研究的必要条件，而目前鲜有文献开展这一工作，本文旨在通过梳理国内外代表性学术文献，总结已有研究工作并加以评析，以期对该领域研究的进一步发展提供助益。

在文献检索与主题选取上，笔者于2019年5月6日采取"数据开放与隐私"与"信息公开与隐私"主题，匹配"精确"模式检索CNKI数据库，另采取"freedom of Information & privacy"与"Open data & privacy"主题检索Web of Science数据库，分别获得中文文献55篇与外文文献148篇，按照主题相关程度分别筛选出39篇中文与36篇外文代表性文献进行深度阅读与分析，并在此基础上将学者们广泛关注与讨论的核心议题确定为综述的主题框架。

## 1　研究的发展历程

政府数据开放中的个人隐私保护研究最早可追溯至1890年隐私保护议题的出现，并与20世纪70年代以来政府信息公开中的个人隐私保护研究一脉相承（见图1）。总体上，可将政府数据开放中个人隐私保护研究的发展历程划分为开始关注、发展准备和重点推进三个阶段。

### 1.1　开始关注阶段（1890－1965）：隐私议题的出现

"隐私"一词最早在1890年被美国著名律师S. Warren和最高法院大法官L. Brandeis定义为"被单独留下的权利"[1]，这也标志着"隐私"作为学术概念正

* 本文系国家社会科学基金青年项目"我国基本公共服务质量监测体系构建研究"（项目编号：18CZZ034）研究成果之一。
作者简介：陈朝兵（ORCID：0000－0003－3318－7871），副教授，博士，硕士生导师，E-mail：chenchaobing818@163.com；郝文强（ORCID：0000－0002－1856－9647），硕士研究生。
收稿日期：2019－07－28　修回日期：2019－10－07　本文起止页码：141－150　本文责任编辑：杜杏叶

ChinaXiv合作期刊

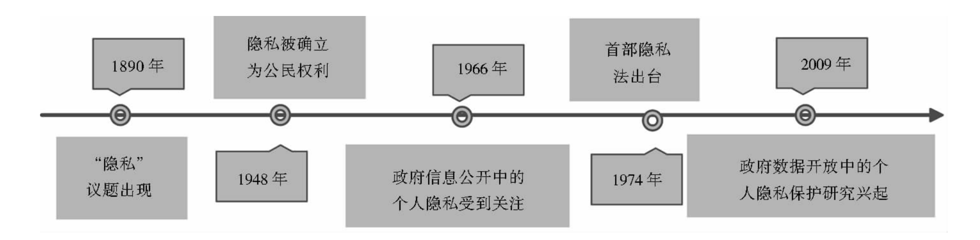



图 1　政府数据开放中个人隐私保护研究的发展历程

式进入学界研究的视野。二战后,个人隐私引起各国重视。联合国安理会在 1948 年缔结的《联合国人权宣言》中宣称:"任何人的隐私、家庭或信件将不受专断的、非法的干涉,也不得非法攻击他人的荣誉和声望;任何人都有受法律保护以反对这种干涉或攻击的权利",把隐私正式确立为公民的一项基本权利,进一步引起了学者们对隐私保护议题的关注[2]。1964 年,美国最高法院审判员 H. Samuel 和学者 H. George 出版《隐私权》一书,系统研究了隐私权利的缘起与范围,提出隐私权可追溯至古老的犹太法律,并援引欧洲和联邦国家的法律案例对隐私权利的范围进行界定[3]。尽管早期的隐私议题研究尚未涉足政府数据开放领域,但其关于隐私概念和隐私权利的讨论无疑为后来政府数据开放中个人隐私保护议题的研究奠定了坚实的理论基础。

### 1.2　发展准备阶段(1966 – 2008):政府信息公开中的个人隐私保护

1966 年,美国颁布《信息自由法》,规定公民享有获得政府信息的权利和行政机关具有向民众提供信息的义务[4]。此后,政府信息公开作为政府行政的发展趋势,逐渐成为隐私风险滋生的土壤。针对此,美国在 1974 年出台《隐私法》,限制行政机构在公开政务信息的过程中泄露个人隐私[5],为政府信息公开中的个人隐私保护提供了法律保障。与此同时,以 E. Volokh、F. H. Cate 等为代表的学者纷纷加入政府信息公开中个人隐私保护研究的行列[6-7]。

在该研究阶段,被置于政府信息公开语境中的个人隐私保护研究不仅对于隐私数据范围的界定比较狭窄,而且研究的话题、视角和方法比较局限。一方面,个人隐私保护的对象被认为是"敏感数据"。1993 年,P. Kumaraguru 和 L. F. Cranor 首次使用"医疗敏感度指数"来描述医疗的隐私担忧程度[8],因此他被认为是最早关注隐私敏感度的学者之一。经过后续学者们的研究,敏感数据被一致认为是指更特别、个人更不愿与他人分享的数据[9],例如"亲密信息"[10]、"高侵入数据"[11]。1995 年,欧盟在《关于涉及个人数据处理的个人保护以及此类数据自由流通的指令(95/46/EC)》中指明敏感数据包含:"种族或民族起源、政治观点、宗教或哲学信仰、工会成员资格、个人医疗或性生活等"[12];另一方面,学者们关注的话题主要包括政府信息公开中个人隐私保护的法律制度、数据管理、公共行政等,学科视角集中在法学、情报学、管理学等社会科学领域,研究方法多采用文献分析法、案例研究法等定性方法[13-14]。例如,D. H. Flaherty 以加拿大不列颠哥伦比亚省《信息自由和隐私保护法》为例,论证了如何通过法律制度设计化解信息公开与隐私保护之间的冲突[15]。

### 1.3　重点推进阶段(2009 至今):政府数据开放中的个人隐私保护

2009 年,美国奥巴马政府发布《开放政府指令》,要求在更大程度、更广范围、更多领域上把政府所掌握的数据向社会开放,由此拉开了世界政府数据开放运动的帷幕。由于"数据"的开放远比"信息"的公开更易引发隐私安全与风险问题,政府数据开放不可避免地给原有的个人隐私保护制度形成剧烈冲击和带来新的挑战。

在政府数据开放新的语境下,该研究阶段扩大了隐私数据的范围,研究话题、学科视角、研究方法等也表现出丰富、广泛、多元的特征。首先,个人隐私保护的对象由敏感数据扩展为所有可识别个人或存在识别个人可能性的一般数据。伴随以大数据、云计算为代表的数据分析技术的出现,即使披露不敏感的、甚至匿名化的个人数据,也可通过重新识别技术侵害到个人隐私[16]。2016 年,欧盟发布的《一般数据保护条例》将个人数据定义为"与已识别或可识别的自然人有关的任何数据",其范围包括但不限于"自然人的姓名、家庭住址、电子信息、身份证件号码、位置信息、IP 地址、浏览记录 IP、手机识别码、医院或医生持有的数据等信息"[17]。其次,学者们在原有研究话题的基础上开始探索政府数据开放中个人隐私保护的匿名技术、数据算法、信息技术等话题,研究的学科视角也实现了法学、情报学、管理学等社会科学与计算机、通信学等自

chinaXiv:202304.00206v1

然科学的结合,并开始使用实验设计、调查研究等研究方法进行实证分析[18-20]。例如,美国北卡大学格林斯堡分校教授 N. Kshetri 通过调查研究,验证了数据的体积、速率、多样性、可变性、复杂性等五个要素对政府数据开放中个人隐私保护的影响[21]。

## 2 政府数据开放中个人隐私保护的概念与缘由

### 2.1 "隐私"概念的嬗变

清晰而准确地界定"隐私"概念是政府数据开放中个人隐私保护研究的逻辑前提。学界对于隐私概念的认识经历了一个从早期充满较大分歧向后期逐渐形成共识的嬗变过程。在该过程中,英美学者是探讨隐私概念的先行者和主力军,贡献了许多卓有见识的理论观点。

21 世纪前,学者们对隐私概念的理解存在价值观念和行为制度两种不同观点。一部分学者持隐私是一种价值观念的观点。比如,英国文学家 S. George 指出,"隐私主要是一种现代的、西方资产阶级的价值观"[22];牛津大学著名哲学家 T. Charles 把隐私界定为西方社会个人对于自我身份概念认知的一种固有特质[23];美国米德维尔伦巴第神学院助理教授 S. Carl 将隐私与羞耻感相结合,认为隐私是保护私人领域免受曝光的羞耻感[24]。与上述学者的观点不同,另一部分学者立足不同学科视角,认为隐私是指人们的行为范式与制度约定。比如,美国哥伦比亚大学政治学系教授 A. F. Westin 早在 1967 年便从法学视角将隐私界定为信息自决权,即"个人、团体或机构要求确定何时、如何以及在多大程度上向他人传达有关他们的信息"[25];美国犹他大学心理学系 I. Altman 从心理学视角将隐私定义为一种个人或者群体自我选择性的控制机制[26];美国南卫理公会大学学者 R. O. Mason 从管理学视角指出隐私问题是信息管理中需要关注的首要伦理问题[27]。

21 世纪以后,随着隐私研究的兴起与渐趋稳定,研究者们对于隐私概念与内涵的理解不断取得共识,普遍认为隐私是一种人类社会的制度产物。正如英国阿伯丁大学研究员 D. J. Solove 所言:"隐私并非自然事物,而是人类基于一系列的范式和法律建立起来的制度。"[28]具体到隐私权利的认识,绝大多数学者认同 A. F. Westin 的观点,认为隐私是指自然人对个人信息的控制权利[29,30]。也有个别学者提出了隐私是群体而非个体权利的新观点,如牛津大学研究员 L. Floridi 认为"隐私权是一种群体权利,是群体作为一个整体拥有的权利,而并非其成员各自拥有的权利"[31]。

### 2.2 政府数据开放中个人隐私保护的缘由

为什么要对政府数据开放中的个人隐私加以保护?围绕这一问题,学者们给出了三种不同解释,即"法定权利说""避免伤害说"和"保障开放说"。

(1)法定权利说。一些学者从法律视角出发,指出保障公民权利是政府数据开放中个人隐私保护的法理依据。但就具体的公民权利内容而言,不同的学者持有不同的观点。有学者认为政府数据开放中个人隐私保护的权利基础是公民隐私权利,如芬兰数据保护委员会秘书 T. Jaatinen 指出,保护个人隐私数据的权利是《欧洲联盟基本权利宪章》规定的一项基本权利,个人隐私数据必须公平、合法地处理[32]。也有学者认为政府数据开放中个人隐私保护的权利基础是言论自由权,如美国学者 B. C. Newell 指出隐私对于确保人们能够在立法和社会规范允许的范围内自由表达自己的意见和采取行动是必要的[33]。还有学者认为政府数据开放中个人隐私保护的权利基础同时包含公民隐私权和言论自由权,如美国学者 D. H. Flaherty 援引加拿大不列颠哥伦比亚省《信息自由和隐私保护法》指出,维护和实现社会公众的隐私和信息自由权益,是公共机构保护个人隐私的双重目标[15]。

(2)避免伤害说。部分学者基于个人隐私保护目标与功能的视角,认为规避隐私伤害是政府数据开放中个人隐私保护的现实需要。在美国斯坦福大学互联网与社会中心高级研究员 R. Calo 看来,政府数据开放导致的隐私伤害可分为主观和客观两个方面:主观伤害体现为个人对自我信息缺乏控制感,客观伤害体现为对一个人意外或胁迫地使用有关该人的信息[34]。郭维嘉和郭少友将社会公众对隐私伤害的隐忧表述为:"如果人们怀疑他们的个人数据会被存储在公共机构,他们可能就不太愿意与公共部门机构联系,特别是对疾病、怀孕、毒品、经济困难或自杀念头有疑问的人,可能不会寻求帮助"[35]。进一步地,加拿大卡尔加里大学教授 T. P. Keenan 指出隐私侵害的不良后果之一是数据的不合理使用和不公平的产生,"公共部门开放的个人数据,可能会被数据经纪人重新使用于新的目的,从而产生不公平"[36]。正是基于以上政府数据开放可能对个人隐私造成的伤害,学者们认为有必要采取措施对政府数据开放中的个人隐私加以保护。

(3)保障开放说。较多学者立足政府数据开放过

chinaXiv:202304.00226v1

ChinaXiv合作期刊

程与效益的角度，认为实施个人隐私保护是保障政府数据开放顺利推行和提高政府数据开放综合效益的必然要求。不难理解，如果政府数据开放侵害到公民个人隐私，将大大损害政府数据开放的好处[37-38]。对此，荷兰海牙安全和司法部研究员 M. S. Bargh 等指出政府数据开放中的个人隐私风险可能会对政府透明度产生收缩性影响，从而导致公民可能对政府失去信任[39]。黄如花和刘龙也指出，个人隐私安全问题关系到公众对政府数据开放的认可度和接受度，如果不能够妥善保护个人隐私安全，政府数据开放将很难进行下去[40]。综上可见，只有切实保护政府数据开放中的公民隐私，才能保障政府数据开放活动的过程顺利和效益最优。

## 3 政府数据开放中个人隐私保护的逻辑悖论与实践困境

学者们注意到，尽管个人隐私保护在政府数据开放中有其价值性和必要性，但其也不可避免地陷入两方面的尴尬境地：一是在逻辑上与数据开放存在冲突与悖论；二是在实践上面临多重困境。

### 3.1 “数据开放”与“隐私保护”的逻辑悖论

从逻辑层面讲，“数据开放”与“隐私保护”之间存在难以调和的张力，意味着要同时实现二者几乎是一件不可能做到的事情。对此，学者们主要从两个层面进行阐释。

首先，数据效用的提高会以数据隐私的牺牲为代价。牛津大学研究员 B. Zevenbergen 等认为，数据效用与数据隐私通常成反比关系，控制隐私泄露风险通常是通过显著降低数据效用来实现的，数据效用的小幅增加则往往需要披露更多的个人信息[41]。荷兰代尔夫特理工大学 M. Andrei 指出，发布原始数据意味着最大的数据效用与没有隐私，不发布数据则意味着完整的个人隐私保护与零数据效用[20]。

其次，数据的组合性和记忆性特征是导致数据开放与个人隐私保护逻辑悖论的根源。荷兰学者 M. S. Bargh 等基于数据的组合特征指出，开放数据从表面看似乎不是个人数据，但是通过将其与其他公开的数据相组合，它可能成为个人数据导致隐私泄露。为了防止隐私泄露，有必要消除隐私敏感属性，但这可会对开放数据的可用性产生负面影响，损坏政府数据开放的效益[39]。田新玲和黄晓芝根据数据的记忆性特征指出，大数据是有记忆的，即使在数据开放过程中及时删除了用户的数据图谱，也不能保证个人数据不被再次利用[42]。

### 3.2 政府数据开放中个人隐私保护的实践困境

沿着“隐私范围－隐私风险－隐私保护”的线索，学者们归纳指出了政府数据开放中个人隐私保护在实践层面的三重困境：隐私范围难界定、隐私风险易发生和隐私保护难实现，并重点阐述了各困境背后的原因。

#### 3.2.1 隐私范围难界定

由于公私数据模糊混杂、数据容易被重新识别、涉及特殊群体等原因，政府数据开放中个人隐私数据范围的界定存在相当困难。田新玲和黄芝晓从数据保存方式的角度分析，认为隐私数据与公共数据模糊混杂在云端，因而哪些属于个人隐私保护的范围难以界定[42]。邹东升从数据识别过程出发，指出个人隐私界定的复杂之处在于：即使最初阶段政府公开的数据中没有个人隐私或者个人信息已被匿名化，但是通过把这些开放的数据与其他数据相联系能够重新识别个人，从而威胁到个人隐私[43]。匈牙利中欧大学 I. Szekely 则以公务员的个人隐私界定为例，指出公务员作为公职人员，其个人隐私与信息公开的界限难以确定，并且公务员的隐私会随着职位的升迁、工作时间的变更而改变[44]。

#### 3.2.2 隐私风险易发生

一部分学者从技术、数据等要素角度分析了政府数据开放中个人隐私风险容易发生的原因。西班牙巴塞罗那自治大学人工智能研究所 G. Navarro 和 V. Torra 指出重新识别技术能够通过多个渠道的信息获取与关联分析，利用政府开放的已经去识别的个人数据重新识别个人[45]。美国北卡大学格林斯堡分校教授 N. Kshetri 认为数据的体积、速率、多样性、可变性、复杂性等均是导致隐私泄露的因素[21]。陈美指出城市政府数据开放中可能导致个人隐私风险的因素包括直接识别数据、间接识别数据、元数据、地址数据、地理坐标、非结构化数据和敏感数据子集[46]。

另一部分学者从数据收集、数据获取、数据使用等过程角度讨论了政府数据开放中个人隐私风险的发生机理。芬兰数据保护委员会秘书 T. Jaatinen 认为数据收集过程中个人自主选择权利的缺失容易导致隐私风险，“通常情况下，数据主体没有真正选择是否向公共部门机构披露其个人数据的权利”[32]。印度国家信息技术学院 S. Gupta 和 P. Kumaraguru 认为数据获取渠道的多样性是造成隐私风险的重要原因[47]。澳大利亚学者 Z. Pingo 和 B. Narayan 认为数据使用过程中信

chinaXiv:202004002v1

息数据被用于已批准的目的以外的其他目的也会导致隐私风险的发生[48]。

### 3.2.3 隐私保护难实现

从制度层面看,学者们认为个人隐私保护制度迟滞与不完善是造成政府数据开放中个人隐私保护难实现的核心原因。具体而言,个人隐私保护制度的迟滞与不完善表现在数据跨域流动缺乏政策规制[49]、个人隐私的概念和范围界定不清晰[50]、个人隐私保护原则不再适用[31]等多个方面。有学者针对我国政府数据开放中个人隐私保护难题指出,“我国开放政府数据的相关利益主体所应享有的权利和责任不明确,个人隐私的概念和范围界定不清晰,这些因素都严重影响了我国个人隐私保护的进程与水平”[51]。

从技术能力层面看,学者们发现技术安全隐患与技术能力不足是造成政府数据开放中个人隐私保护困境的重要原因。美国佛罗里达州立大学研究员 P. T. Jaeger 等从信源、信道、信宿三个方面论证了信息技术进步使得个人隐私保护比以前更加困难。在信源方面,技术发展为收集公众信息提供了更为便捷的手段;在信道方面,技术让信息的传播更为便捷;在信宿方面,技术让信息利用方式和利用主体比以前更为复杂[52]。美国北卡大学格林斯堡分校教授 N. Kshetri 指出政府可能缺乏在高峰数据流量期间安全存储大量数据和管理非结构化数据的技术能力,从而造成隐私泄露风险[21]。

从心理与文化层面看,学者们认为公众个体的隐私观念、隐私披露心理、隐私侵害防范能力等在很大程度上影响政府数据开放中个人隐私保护的实现。荷兰奈梅亨大学计算与信息科学研究所 B. Jacobs 认为隐私保护意识缺乏是阻碍个人隐私保护的重要原因。他指出,社交网络服务为人们提供了在网上被看到和实时分享经验的途径,而对许多人来说,“被看到的愿望比保护私人信息的愿望更强烈”[53]。德国霍恩海姆大学媒体心理学系 S. Trepte 等认为引致隐私风险发生的一种主要原因就是公众缺乏对隐私保护工具的重视心理、基本认知和正确应用[54]。

## 4 政府数据开放中个人隐私保护的实现路径

理论研究的最终目的是科学指导实践。对实现路径的研究直接关系到政府数据开放中个人隐私保护实践的指导,因而受到学界的广泛关注。概括而言,学者们主要基于制度、组织、技术和素养四个层面探讨了政府数据开放中个人隐私保护的实现路径。

### 4.1 强化个人隐私保护制度设计

制度设计是政府数据开放中个人隐私保护首当其冲的环节,对此学者们侧重围绕法律制定、制度内容与制度工具展开讨论。

在法律制定方面,学者们大致提供了三种思路:一是制定国家层面的个人隐私保护法律。比如,张晓娟等认为我国应当制定《个人信息保护法》,完善政府数据开放和个人隐私保护的法律体系[55]。二是在制定国家个人隐私保护法律的同时,还需要国际性的个人隐私保护法律。比如,全球言论自由运动高级法律顾问 D. Banisar 以尼日利亚政府数据开放中的个人隐私保护为例,指出该国既有《尼日利亚宪法》规定公民的隐私权,保护他们的家园、通信、电话交谈和电报通信,也有类似于《非洲儿童权利与福利宪章》的国际法律保护儿童的隐私权利[56]。三是将信息数据自由与个人隐私保护二者进行联合立法。比如,匈牙利中欧大学 I. Szekely 指出信息隐私和信息自由是相互关联、相互依存的概念,可以通过这两个领域的联合或至少相互关联的立法模式来实现政府数据开放中个人隐私保护的目标[44]。

在制度内容方面,学者们提出了遵循个人隐私保护原则、实施隐私数据审查、完善网站隐私政策等观点。比如,荷兰海牙安全和司法部 M. S. Bargh 等以《荷兰隐私保护法》为例,指出个人隐私保护的原则包括终结原则、合法原则、相称原则、辅助原则、透明度原则和权利原则[39]。又如,黄如花和刘龙提出我国政府数据开放中个人隐私保护的政策内容应包括严格的数据审查,制定统一的数据脱敏处理标准规范,提升数据开放的标准化、规范化,尽可能减少隐私风险发生的可能性[41]。再如,冯昌杨以我国政府数据开放门户为例,要求制定较为全面的网站隐私政策,需要包含用户信息的内容、目的和方法、信息的收集、个人信息保护、在线评论、浏览信息收集、网站安全、外部链接、免责声明以及政策调整等内容[57]。

在制度工具方面,学者们针对隐私影响评估这一政策工具,从评估的内容、方法与实施方面进行了研究。关于隐私影响评估的内容,印度国家信息技术学院 S. Gupta 和 P. Kumaraguru 认为包括伤害程度、可重复性、可利用性、影响程度、可发现性等多个维度[47]。关于隐私影响评估的方法,荷兰学者 M. Andrei 认为可采用阈值规则、样本权重、使用启发式与使用记录链

chinaXiv:202301.00026v1

ChinaXiv合作期刊

接等[20]。关于隐私影响评估的实施，黄如花和刘龙认为应当建立整个数据生命周期的隐私影响分析机制，在政府数据收集、处理、组织、开放和使用各个阶段加强隐私影响的研判和审查[40]。

## 4.2　设置个人隐私保护组织机构

学者们虽然一致认为应把设置组织机构作为政府数据开放中个人隐私保护的实现路径，但在是否设置独立的个人隐私保护机构以及政府部门间是合作还是独立关系上却存在较大争议。

关于设置个人隐私保护机构，大多数学者认为应当建立专门的个人隐私保护机构。比如，美国佛罗里达州立大学 P. T. Jaeger 等提出在政府数据开放的背景下，政府应该采用中央集权的管理机构来保护个人隐私[52]。黄如花和刘龙也认为我国应设立个人隐私保护的专门机构。该机构在法律地位上是独立的，不受一般政府部门的制约，仅需向某一特定机构负责即可[58]。与上述意见不同，个别学者认为现阶段我国设立独立的个人隐私保护机构并不现实，可以考虑将个人隐私保护的职责交付既有的开放数据安全管理机构[47]。

关于政府部门之间的关系，大部分学者认为应当建立跨职能部门的政府数据开放个人隐私保护机制。例如，有学者认为同一层级政府的不同部门和机构之间的合作与沟通，有助于提高政府数据开放中个人隐私保护的效果[59]。不过，也有学者认为在政府数据开放活动中进行数据集中管理存在安全隐患，应该重创信息孤岛。比如，荷兰代尔夫特理工大学技术政策与管理学院 M. Janssen 和 J. Hoven 认为过去的努力主要集中在拆除孤岛上，而如今应该重新创造信息孤岛，以此防止信息逐渐集中，建立保护个人隐私的机制[60]。

## 4.3　应用个人隐私保护先进技术

在政府数据开放中个人隐私保护的技术路径研究中，匿名技术是一种广受学者青睐的技术类型。比如，以英国开放大学研究助理 C. Perera 等为代表的学者提出了诸多可应用于政府数据开放中个人隐私保护的匿名策略，包括最小化设计、假名、接受透明度与公开性、隐藏、分离、聚合、控制等[61]。

然而，随着大数据时代背景下政府信息公开向政府数据开放的转型与深入发展，数据重新识别可能性的增大使得匿名技术可能不足以达到个人隐私保护的目的，因而不少学者开始进行信息熵、区块链以及其他新兴技术在政府数据开放中个人隐私保护的应用研究。比如，韩国大田高等科学技术学院 S. H. Kim 等通过实验设计，提出了信息熵的个人隐私保护模型。该模型以 S 属性集作为一个基本单位来度量组织的熵，并且可以从组织的所有风险中获得数据的熵值，在精细单元的基础上实现更准确的重新识别风险，从而保护隐私并保证数据的可用性[19]。又如，美国学者 F. Bonomi 等提出了雾计算的个人隐私保护技术。他指出，雾计算是一种将云计算和服务扩展到网络边缘的范例，其独特的雾特征是其接近最终用户、密集的地理分布以及对移动性的支持，具备减少潜在侵犯隐私行为的作用[62]。

## 4.4　培育公民隐私保护素养

实现政府数据开放中的个人隐私保护并非政府主体一方之责和一方之事，而是也有赖于个人隐私保护的客体——公民一方的参与和配合。其中，具有相当重要性的是公民的隐私保护素养。正如邓胜利和王子叶指出："隐私保护素养是影响政府数据开放隐私保护的关键因素。提高公民隐私保护素养对于政府数据开放隐私保护具有重要作用。"[63]

围绕政府数据开放中的公民隐私保护素养议题，学者们重点讨论了培育公民隐私保护素养的途径。在部分学者看来，提高公民隐私保护素养主要依靠教育与培训。比如，澳大利亚学者 Z. Pingo 与 B. Narayan 认为教育对于帮助公民在为各种目的公布个人信息时做出合理和主动的决定非常重要，这也能帮助公民在隐私问题上做出有效决策[48]。另一部分学者则认为，公民自主学习才是提高公民隐私保护素养的主要途径。比如，以色列学者 M. Weinberger 等指出现在大部分在线隐私保护工具的信息都是公开的，而且这些工具很容易学习和使用，用户要自觉地发掘并学习如何保护自己个人隐私的方法与技能[64]。

除了上述制度、组织、技术和素养四种路径外，一些学者还提出了设置个人隐私、保护群体性隐私、利益相关者集体行动等实现政府数据开放中个人隐私保护的观点。比如，荷兰阿姆斯特丹大学信息法律研究所研究员 D. Sloot 提出有必要引入有关第三方重新使用公民数据的个人隐私设置，让每个人都向政府登记自己的隐私设置，并由公民自主选择数据发布对象、范围、使用目的等，以通过向公民提供第三方重新使用其个人数据获得的一定比例的利润来激励他们加入[65]。又如，牛津大学研究员 L. Floridi 认为应当通过保护群体隐私来实现政府数据开放中个体隐私的保护，这就好比"个别沙丁鱼可能认为，渔网试图抓住它。事实上，渔网试图抓住整个浅滩。因此，如果要拯救沙丁

chinaXiv:202304.00260v1

鱼,需要保护的是浅滩。有时,保护个人的唯一办法是保护个人所属的群体"[31]。再如,牛津大学牛津互联网研究所 N. Kshetri 从利益相关者的角度提出,"政府数据开放隐私保护过程中,政府必须采取合理的监管措施,制定更高的隐私标准;企业应积极主动地制定风险管理体系并深入了解风险的根源;公民则可以通过维权对政府施加压力要求其合法合理地开放和使用数据"[21]。

## 5 研究总结与未来展望

综上可以看出,个人隐私保护是国内外学界在很长一段时间里保持密切和持续关注的热点议题,特别是在政府数据开放这一新的背景下引起了更多学者的研究和重视。其中,荷兰代尔夫特大学技术政策与管理学院 M. Andrei 和 M. Janssen、英国牛津大学牛津互联网研究所 L. Floridi 和 B. Zevenbergen、美国佛罗里达州立大学信息研究学院 P. T. Jaeger、中国武汉大学信息管理学院黄如花等是该领域具有标识度的研究力量。已有研究围绕政府数据开放中个人隐私保护的概念、缘由、逻辑悖论、实践困境、实现路径等做出了集中讨论,不仅取得了比较丰硕的研究成果,而且初步形成了学科多元、方法多样的政府数据开放中个人隐私保护"是什么""为什么""怎么样""怎么办"的专题研究体系,具体如图 2 所示:

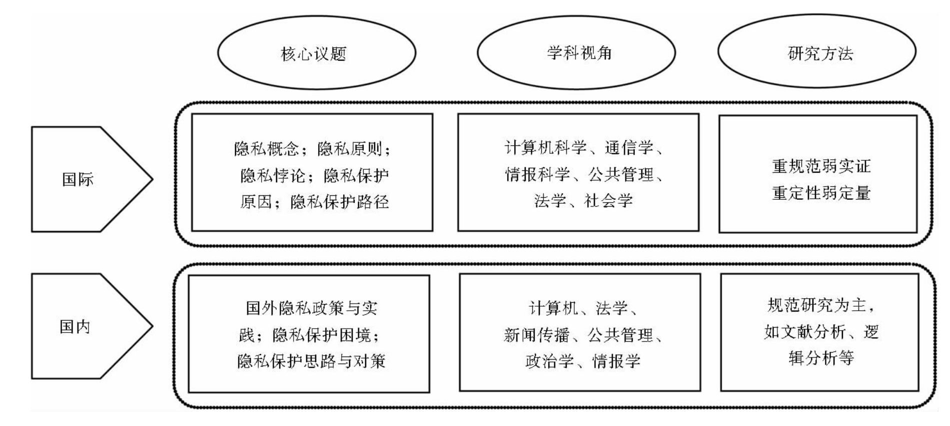



图 2 国内外政府数据开放中的个人隐私保护研究图谱

整体上,已有研究呈现出三个特征:一是研究时段的转化性与延续性。该领域研究经历了政府信息公开和政府数据开放两个时段,且两个时段之间具有内在的转化性和延续性。二是研究视角的跨学科性。不仅得到情报学、法学、政治学、公共管理学等社会学科的讨论,而且也受到了计算机、通讯学等自然学科的关注。三是研究观点的共识性与分歧性并存。如在隐私保护缘由、隐私保护与数据开放间的张力、隐私保护困境与路径等方面取得了基本共识,而在隐私的概念与内涵、隐私保护的权利基础、隐私保护机构设置等问题上仍存较大争议。

已有研究尚存在一定的局限与不足之处,表现在:在研究议题上,对于隐私基本问题(如隐私的含义、特质、范围、发生条件)[66]的讨论尚不充分,同时对于个人隐私保护问题的挖掘也比较有限;在研究对象上,对于政府数据开放中个人隐私保护这一研究对象的系统性和复杂性的把握还不深入,未能有效运用系统思维和复杂思维加以应对;在研究工具和视角上,对于相关学科理论工具的创新性引入比较缺乏,研究视角也未超越结果视角和静态视角;在研究方法上,尚停留在规范分析层面,开展的实证研究明显偏少,不能满足科学研究的需要。

站在政府数据开放这一新兴时代背景下,着力于当前政府数据开放中个人隐私保护的理论指导需要[67],同时以更宏大、全面和整体的研究视野加以审视,我们认为未来国内该领域的研究应当重点强化和突破以下四个方面:

(1)进一步拓展个人隐私保护的研究问题范畴,同时聚焦个人隐私保护中的关键问题与重大问题研究。下一步的研究工作需要拓展政府数据开放中个人隐私保护的相关研究问题,包括隐私的含义、特质与本质,政府数据开放中个人隐私的边界与范围,政府数据开放中个人隐私风险的发生机理、演化过程与影响因素等。除此外,值得挖掘的政府数据开放中个人隐私保护的研究问题还包括隐私的影响、危害及其评估,隐私风险与问题的分类,个人隐私保护中的价值平衡与冲突管理,隐私管理与保护的典型经验和教训、隐私保护的体制机制设计与系统化操作策略,等等。最后,针

chinaXiv:202304.00260v1

ChinaXiv合作期刊

对已研究和上述列举的诸多问题，需要进一步识别其中的关键问题和重大问题，如隐私问题的发生条件与机理，隐私保护的系统化策略等，继而投入大量的研究精力和资源加以突破，实现研究的价值聚焦。

（2）把握政府数据开放中个人隐私保护问题的复杂性，并注重系统思维在研究中的运用。政府数据开放的公共活动性质决定了其中的个人隐私保护问题具有复杂性，表现在政府数据开放中个人隐私保护的价值取向和目标定位不是单一的和一致的（如存在隐私保护与透明开放两种具有张力的价值取向与目标定位），以及牵涉的要素是综合的和复杂的（如涉及政府体制、公共政策、社会需求、公共权力、政府责任等多种要素）。为此，面对政府数据开放中的个人隐私保护问题，研究者必须综合运用情报学、政治学、行政学、管理学、法学等多学科知识，同时运用系统、复杂、整体和联系思维加强对政府数据开放中个人隐私保护问题的把握。比如，对于如何实施政府数据开放中个人隐私保护这一问题，在主体上需要同时考虑政府、企业、社会组织、专业机构、新闻媒体、公民个体等，在机制上需要同时考虑数据管理机制、行政法律规制机制、行政伦理引导机制、政府责任约束机制等。

（3）丰富和创新相关理论工具的应用，突出研究的过程视角和动态视角。比如，隐私问题产生于政府数据开放的全生命周期中，因而数据生命周期理论无疑可为政府数据开放中个人隐私保护的分析提供有效的分析框架，即可依次从数据生命周期的各个阶段深入分析隐私风险的产生因素并进而提出防范化解措施。又如，政府数据开放中个人隐私保护问题实际是政府、企业、公民三方主要利害关系人之间相互博弈的过程，因而利益相关者理论和博弈理论均可提供适当的理论基础，并为隐私风险产生、隐私保护实现等问题的解释和解决提供新的视角。在丰富和创新相关理论工具应用的同时，该领域还应当更多转向和推进过程视角和动态视角的研究，以弥补已有研究更多偏向于结果视角和静态视角的研究不足。比如，加强政府数据开放中个人隐私风险产生的诱因及其形成机理的过程研究，加强政府数据开放中个人隐私管理与保护机制的动态研究等。

（4）引入历史分析法、比较分析法等分析方法，同时加强案例研究等实证研究的开展。隐私保护是一个由来已久的问题，其在较长的历史时期中蕴藏着一定的经验与教训，以及特定的演化与发展规律。同时，隐私保护是不同国家、不同地域、不同层级的政府面临的共同问题，相应地，隐私保护在这些不同政府之间必然积累了充满共性和差异的若干实践。这意味着，历史分析法和比较分析法对于现今政府数据开放中的个人隐私保护研究具有很好的适用性，应当将其作为目前该领域规范分析的重要方法加以引入和补充。又由于个人隐私保护具有很强的实践性，政府数据开放中个人隐私保护的实证研究可选取案例研究作为重要的实证研究方法，以改变该领域重规范轻实证的研究现状。比如，可通过探寻国内外政府数据开放中个人隐私保护的典型个案，为政府数据开放中个人隐私保护理论命题提供经验支撑。

chinaXiv:202304.00260v1

**参考文献：**

[1] WARREN S, BRANDEIS L. The right to privacy[J]. Harvard law rev, 1890, 4(5): 193-220.

[2] The universal nations declaration of human rights[EB/OL]. [2019-05-06]. https://www.un.org/en/universal-declaration-human-rights/index.html.

[3] HOFSTADTER S H, HOROWITZ G. The right of privacy: with a reprint of "Right to Privacy" by Samuel D. Warren and Louis D. Brandeis[M]. New York: Central Book Company, 1964.

[4] What is FOIA? [EB/OL]. [2019-05-06]. https://www.foia.gov/about.html.

[5] The privacy act of 1974 [EB/OL]. [2019-05-06]. https://www.archives.gov/about/laws/privacy-act-1974.html.

[6] VOLOKH E. Freedom of speech and information privacy: the troubling implications of a right to stop people from speaking about you [J]. Stanford law review, 2000, 52(5): 1049-1124.

[7] CATE F H, FIELDS D A, MCBAIN J K. The right to privacy and the public's right to know: the "central purpose" of the freedom of information act[J]. Administrative law review, 1994, 46(1): 41-74.

[8] KUMARAGURU P, CRANOR L F. Privacy indexes: a survey of Westin's studies[EB/OL]. [2019-05-06]. http://www.cs.cmu.edu/~ponguru/CMU-ISRI-05-138.pdf.

[9] 王敏. 价值趋同与文化存异：中美"千禧一代"大学生对敏感数据的感知对比[J]. 新闻与传播评论, 2018, 71(2): 28-41.

[10] NISSENBAUM H. Protecting privacy in an information age: the problem of privacy in public[J]. Law and philosophy, 1998, 17 (5): 559-596.

[11] PESCIOTTA D T. I'm not dead yet: katz, jones, and the fourth amendment in the 21st century[J]. The case western reserve law review, 2012, 63(2): 187-255.

[12] Data protection directive, article 4[EB/OL]. [2019-05-06]. http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf.

[13] COMEAU P A, OUIMET A. Freedom of information and privacy: Quebec's innovative role in north america[J]. IRB, 1995, 23(4): 13-15.

chinaXiv:202304.00260v1

[14] SHEINKOPF C M. Balancing free speech, privacy and open government: why government should not restrict the truthful reporting of public record information[J]. Ucla L. rev, 1997,44(5):1567－1611.

[15] FLAHERTY D H. Balancing open government and privacy protection [EB/OL]. [2019－05－06]. http://www.austlii.edu.au/au/journals/PrivLawPRpr/1999/56.html.

[16] MACHANAVAJJHALA A, REITER J P. Big privacy: protecting confidentiality in big data[J]. Xrds crossroads the acm magazine for students, 2012, 19(1):20－23.

[17] The general data protection regulation[EB/OL]. [2019－05－06]. http://www.consilium.europa.eu/en/policies/data－protection－reform/data－protection－regulation/.

[18] 张文亮. 个人数据保护立法的要义与进路[J]. 江西社会科学,2018,38(6):169－176.

[19] KIM S H, JUNG C, LEE Y J. An entropy－based analytic model for the privacy－preserving in open data[EB/OL]. https://www.onacademic.com/detail/journal_1000039815718510_95d6.html.

[20] ANDREI M. Publishing privacy sensitive open data[EB/OL]. [2019－05－06]. http://insy.ewi.tudelft.nl/sites/default/files/thesis_final.pdf.

[21] KSHETRI N. Big data's impact on privacy, security and consumer welfare[M]. Oxford:Pergamon Press, 2014.

[22] STEINER G. Literature and post－historicism[M]. London: Cato Winders Company, 1971.

[23] TAYLOR C. The crisis of modernity[M]. Cambridge: Cambridge University Press, 1985.

[24] SCHNEIDER C D. Shame, exposure, and privacy[M]. Cambridge: Cambridge University Press, 2000.

[25] WESTIN A F. Science, privacy, and freedom: issues and proposals for the 1970s. Part I——the current impact of surveillance on privacy[J]. Columbia law review, 1966, 66(6):1003－1050.

[26] ALTMAN I. Privacy regulation: Culturally universal or culturally specific? [J]. Journal of social issues,1977,33(3): 66－84.

[27] MASON R O. Four ethical issues of the information age[J]. MIS quarterly,1986,10(1): 5－12.

[28] SOLOVE D J. Understanding privacy[J]. Social science electronic publishing, 2008, 59(7):57－58.

[29] GÜRSES F S. Multilateral privacy requirements analysis in online social network services[J]. Journal of vibration & shock,2010,27(9): 139－141.

[30] GUTWIRTH S, POULLET Y, HERT P D, et al. Reinventing data protection? [J]. Identity in the information society, 2010, 26(6):673－676.

[31] FLORIDI L. Open data, data protection, and group privacy[J]. Philosophy & technology, 2014,27(1):1－3.

[32] JAATINEN T. The relationship between open data initiatives, privacy, and government transparency: a love triangle? [J]. International data privacy law, 2016, 6(1):28－38.

[33] NEWELL B C. Technopolicing, surveillance, and citizen oversight: a neorepublican theory of liberty and information control[J]. Government information quarterly, 2014,31(3):421－431.

[34] CALO R. The boundaries of privacy harm[J]. Social science electronic publishing, 2011,86(3):1131－1162.

[35] 郭维嘉,郭少友. 开放政府数据中的个人隐私风险类型及消解机制研究[J]. 河南科技,2018(4):21－24.

[36] KEENAN T P. Are they making our privates public? － Emerging risks of governmental open data initiatives[C]//IFIP primelife international summer school on privacy and identity management for life. Berlin: Springer, 2011:1－13.

[37] JANSSEN M, CHARALABIDIS Y, ZUIDERWIJK A. Benefits, adoption barriers and myths of open data and open government[J]. Information systems management, 2012, 29(4):258－268.

[38] RONALD M, PETER C, SUNIL C. Reconciling contradictions of open data regarding transparency, privacy, security and trust[J]. Journal of theoretical and applied electronic commerce research, 2014,9(3):32－44.

[39] BARGH M S, CHOENNI S, MEIJER R. On addressing privacy in disseminating judicial data: towards a methodology[J]. Transforming government: people, process and policy. 2017,11(1): 9－41.

[40] 黄如花,刘龙. 我国政府数据开放中的个人隐私保护问题与对策[J]. 图书馆, 2017(10):1－5.

[41] ZEVENBERGEN B, BROWN I, WRIGHT J, et al. Ethical privacy guidelines for mobile connectivity measurements FINALv2.docx. [EB/OL]. [2019－05－06]. http://dataethics.github.io/proceedings/EthicalPrivacyGuidelinesforMobileConnectivityMeasurements.pdf.

[42] 田新玲,黄芝晓. "公共数据开放"与"个人隐私保护"的悖论[J]. 新闻大学, 2014(6):55－61.

[43] 邹东升. 政府开放数据和个人隐私保护:加拿大的例证[J]. 中国行政管理, 2018(6):75－82.

[44] SZEKELY I. Freedom of information versus privacy: friends or foes? [M]. Berlin: Springer Netherlands, 2009.

[45] NAVARRO－ARRIBAS G, TORRA V. Data privacy: a survey of results[J]. Studies in computational intelligence. 2015,10(3):27－37.

[46] 陈美. 城市政府开放数据的隐私风险及其技术控制策略[J]. 图书馆建设,2018,290(8):16－21,27.

[47] GUPTA S, KUMARAGURU P. OCEAN: Open－source collation of e government data and networks － understanding privacy leaks in open government data[EB/OL]. [2019－05－06]. https://arxiv.org/pdf/1312.2784.pdf.

[48] PINGO Z, NARAYAN B. When personal data becomes open data: an exploration of lifelogging, user privacy, and implications for privacy literacy[EB/OL]. [2019－05－06]. https://link.springer.com/content/pdf/10.1007%2F978－3－319－49304－6_1.pdf.

[49] BAUMER D L, EARP J B, POINDEXTER J C. Internet privacy law: a comparison between the United States and the European Union[J]. Computers & security, 2011, 23(5):400－412.

[50] 陈美. 德国开放政府数据中的个人隐私保护研究[J]. 图书馆, 2018(8):11－16.

ChinaXiv合作期刊

[51] 黄如花,李楠.美国开放政府数据中的个人隐私保护研究[J].图书馆,2017(6):19-24,76.

[52] JAEGER P T, MCCLURE C R, FRASER B T. The structures of centralized governmental privacy protection: approaches, models, and analysis[J]. Government information quarterly,2002, 19(3):317-336.

[53] JACOBS B. Two of the grand changes through computer and network technology[M]. Berlin:Springer, 2013.

[54] TREPTE S,TEUTSCH D,MASUR P K,et al. Do people know about privacy and data protection strategies? towards the "online privacy literacy scale"(OPLIS) [M]. Berlin:Springer Netherlands, 2015.

[55] 张晓娟,王文强,唐长乐.中美政府数据开放和个人隐私保护的政策法规研究[J].情报理论与实践, 2016,39(1):38-43.

[56] BANISAR D. Linking ICTs, the right to privacy, freedom of expression and access to information[J]. Social science electronic publishing, 2010,16(1):124-154.

[57] 冯昌扬.政府开放数据门户网站隐私政策比较研究[J].数字图书馆论坛, 2016(7):52-56.

[58] 黄如花,刘龙.英国政府数据开放中的个人隐私保护研究[J].图书馆建设, 2016(12):47-52.

[59] Strategies for protecting privacy in open data and proactive disclosure | Scassa | Canadian journal of law and technology[EB/OL].[2019-05-06]. https://ojs.library.dal.ca/CJLT/article/view/8480/7286.

[60] JANSSEN M, HOVEN J V D. Big and open linked data (BOLD) in government: a challenge to transparency and privacy? [J]. Government information quarterly, 2015, 32(4):363-368.

[61] PERERA C, RANJAN R, WANG L. End-to-end privacy for open big data markets[J]. IEEE cloud computing, 2015, 2(4):44-53.

[62] BONOMI F, MILITO R, ZHU J, et al. Fog computing and its role in the internet of things[EB/OL].[2019-05-06]. http://conferences.sigcomm.org/sigcomm/2012/paper/mcc/p13.pdf.

[63] 邓胜利,王子叶.国外在线隐私素养研究综述[J].数字图书馆论坛, 2018(9):66-72.

[64] WEINBERGER M, ZHITOMIRSKY-GEFFET M, BOUHNIK D. Factors affecting users online privacy literacy among students in Israel[J]. Online information review,2017,41(1):655-671.

[65] SLOOT B V D. On the fabrication of sausages, or of open government and private data[J]. Social science electronic publishing, 2013,3(2):136-151.

[66] 李延舜.大数据时代信息隐私的保护问题研究[J].河南社会科学, 2017,25(4):67-73,124.

[67] 侯水平.大数据时代数据信息收集的法律规制[J].党政研究,2018(2):22-28.

**作者贡献说明:**

陈朝兵:设计论文框架,提出研究方案,负责论文撰写与修改;

郝文强:收集、梳理文献,撰写论文初稿,参与论文后期修改。

chinaXiv:202304.00260v1

## Review of Domestic and Foreign Research on the Personal Privacy Protection in the Open Government Data

Chen Chaobing Hao Wenqiang

School of Public Administration,Southwest University of Finance and Economics,Chengdu 611130

Abstract: [**Purpose/significance**] The research literature of personal privacy protection in the opening of government data at home and abroad is systematically combed, aiming to provide reference and reference for further research in this field in China. [**Method/process**] Using the method of literature research, induction, comparison and synthesis, this paper made a review of the concept, reason, logical paradox, practical dilemma and realization path of individual privacy protection in the opening of government data. [**Result/conclusion**] The results show that the research in this field has experienced three stages: the emergence of privacy issues, the protection of personal privacy in government information disclosure and the privacy protection of individuals in government data opening. The research on related topics has achieved fruitful results, the characteristics of interdisciplinary research are obvious, and the research points of view are both consensus and differences. It is suggested that the domestic research should be carried out from four aspects: expanding the scope of research topics, focusing on key issues, grasping the complexity of research issues, paying attention to applied systematic thinking, enriching the application of theoretical tools, highlighting the process and dynamics of research, introducing historical, comparative and other analytical methods to strengthen empirical research.

Keywords: open government data government information disclosure personal privacy privacy protection documentary review